Session 5 Overview

Microprocessors

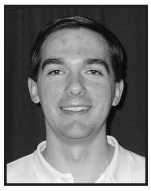
Chair: *Stefan Rusu, Intel, Santa Clara, CA*

Associate Chair: *Jim Warnock, IBM, Yorktown Heights, NY*

This year's microprocessor session highlights the accelerating trend towards multi-core integration on a single chip, setting new records for the number of threads supported. This year, most of the chips presented feature at least 4 cores, a marked increase in integration from last year. Cache sizes continue to increase, with typically 2MB to 4MB L2/L3 integrated on-die. These advances in integration are enabled by ever increasing numbers of interconnect layers, with two papers describing the use of 11 interconnect layers, and also the progression to 65nm technology. Power continues to be a critical concern, as all papers describe special active power and leakage reduction techniques. Most of the designs focus on low-voltage operation, presenting circuit techniques for minimum supply-voltage reduction, and two papers describe independent dynamic frequency control of individual cores. Even as the industry shifts towards multi-core, power-constrained designs, frequency improvements are still possible, with a new record being reported this year for microprocessor operating frequency.

The session starts with Paper 5.1 from IBM describing the dual-core POWER6™ design, operating at frequencies up to 5GHz. The 341mm^2 die contains over 700M transistors and is fabricated in 65nm SOI technology with 10 levels of interconnect. Paper 5.2 from Intel is also a high-frequency design, describing a network-on-chip architecture that contains 80 tiles including floating-point cores and packet-switched routers, all operating at 4GHz. The 100M transistor die, implemented in 65nm technology, achieves 1TFLOP peak performance using a variety of circuit techniques to keep power dissipation under 100W. The third paper, 5.3, from Renesas, Hitachi, and Waseda U describes a quad-core SoC implemented in 90nm triple-V_T CMOS technology with 8 levels of metal. This chip also employs advanced techniques to reduce average operating power with the capability of operating each processor core dynamically at a different frequency. With a die size of less than 100mm^2 this chip achieves a performance 4320MIPs and a floating-point performance of 16.8GFLOPS.

In the second half of the session, all four papers describe multi-core microprocessors implemented in 65nm CMOS technology. The first quad-core Opteron™ processor is presented by AMD in Paper 5.4. Similar to the previous paper, this design allows independent dynamic frequency control of each core. The 450M transistor chip is built in 65nm SOI technology with 11 interconnect levels. Each core includes 512kB of L2 cache, with a 2MB L3 cache shared by all cores. Paper 5.5 from PA Semi describes a power-efficient SoC integrating two 2GHz Power™ cores with shared 2MB L2 cache, coherent crossbar interconnect, and memory and I/O subsystem. The 115mm^2 die is implemented in a 65nm 8-layer metal technology. Paper 5.6 from Intel describes the dual-core, and quad-core (on MCM) implementations of the Core™ architecture. The 143mm^2 die has 291M transistors in an 8-layer metal 65nm process. The shared 4MB L2 cache uses PMOS power gating to minimize leakage. The processor has a wide operating range, from 0.85V to 1.325V, at frequencies from 1 to 3GHz. Finally, Paper 5.7 from Sun Microsystems presents the 8-core, 64-thread second generation Niagara SPARC™ SoC, doubling the number of threads from the previous design. The chip integrates 4MB L2 cache, one x8 PCI-express, two 10G Ethernet ports, and 8 FBDIMM ports. The 500M transistor chip, with a die size of 342mm^2, is implemented in an 11-metal 65nm triple-V_T CMOS process.



5.1 Design of the POWER6™ Microprocessor **1:30 PM**

J. Friedrich, IBM, Austin, TX

The POWER6™ microprocessor combines ultra-high frequency operation, aggressive power reduction, a highly scalable memory subsystem, and mainframe-like reliability, availability, and serviceability. The 341mm^2 700M transistor dual-core microprocessor is fabricated in a 65nm SOI process with 10 levels of low-k copper interconnect. It operates at clock frequencies over 5GHz in high-performance applications, and consumes under 100W in power-sensitive applications.

5.2 An 80-Tile 1.28TFLOPS Network-on-Chip in 65nm CMOS **2:00 PM**

S. Vangal, Intel, Hillsboro, OR

A 275mm^2 network-on-chip architecture contains 80 tiles arranged as a 10×8 2D array of floating-point cores and packet-switched routers, operating at 4GHz. The 15-FO4 design employs mesochronous clocking, fine-grained clock gating, dynamic sleep transistors, and body-bias techniques. The 65nm 100M transistor die is designed to achieve a peak performance of 1.0TFLOPS at 1V while dissipating 98W.

5.3 A 4320MIPS Four-Processor Core SMP/AMP with Individually Managed Clock Frequency for Low Power Consumption

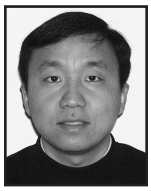
Y. Yoshida, Renesas Technology, Tokyo, Japan **2:30 PM**

A 4320MIPS four-core SoC that supports both SMP and AMP for embedded applications is designed in 90nm CMOS. Each processor-core can be operated with a different frequency dynamically including clock stop, while keeping data cache coherency, to maintain maximum processing performance and to reduce average operating power. The 97.6mm^2 die achieves a floating-point performance of 16.8GFLOPS.

5.4 An Integrated Quad-Core Opteron™ Processor **3:15 PM**

S. Searles, AMD, Austin, TX

An integrated quad-core ×86 processor is implemented in a 65nm 11M SOI CMOS process. Based on an enhanced Opteron™ core, the SoC-developed processor employs power- and thermal-management techniques throughout the design. The SRAM cache designs target process variation considerations and future process scalability. A DDR2/DDR3 combo-PHY and HT3 I/Os provide high-bandwidth interfaces.



5.5 A 25W SoC with Dual 2GHz Power™ Cores and Integrated Memory and I/O Subsystems **3:45 PM**

Z. Chen, PA Semi, Santa Clara, CA

An SoC is presented with dual 2GHz Power™ cores, coherent crossbar interconnect, 2MB L2 cache, and memory and I/O subsystem. The chip consumes a maximum of 25W of power. The 115mm^2 die is implemented in a 65nm 8M process with low-power design techniques. Circuits to improve system performance under power constraints are discussed.

5.6 Implementation of the 65nm Dual-Core 64b Merom Processor **4:15 PM**

N. Sakran, Intel, Haifa, Israel

Merom is a dual-core 64b processor implementing the Core™ architecture. The 143mm^2 die has 291M transistors in a 65nm 8M process. The shared 4MB 16-way L2 cache uses PMOS power gating to minimize leakage. The processor operates in a wide core frequency range of 1 to 3GHz, a bus frequency range of 666 to 1333MHz and voltage range of 0.85 to 1.325V, while providing 40% better power performance.

5.7 An 8-Core 64-Thread 64b Power-Efficient SPARC SoC **4:45 PM**

U. Nawathe, Sun Microsystems, Sunnyvale, CA

The 8-core 64-thread 64b power-efficient 2nd-generation Niagara SPARC SoC has 4MB L2 cache with one ×8 PCI-Express, two 10G Ethernet (XAUI), and 8 FBDIMM ports. The on-chip SerDes provide greater than 1Tb/s bandwidth. The 500M transistor chip with a die size of 342mm^2 is implemented in a 11M 65nm triple-V_t CMOS process.